

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2003, 20:41:35 ; Search time 2962 Seconds

(without alignments)
4834.021 Million cell updates/sec

Title: US-09-745-506-37

Perfect score: 350

Sequence: 1 MDLKAISLNDFAISFAE.....LEKNITILSETRDELQYV 350

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.p2n.model -DEV=xlh
-Q/cg2_1/USPTO.spool/US09745506/runatc.22082003_104427_7264/ap2_query.fasta.1.519
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2nol1.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-BITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptio
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09745506.qcgn.1.1.2372.etruncat.22082003_104427_7264 -NCPU=6 -ICPU=3
-NO_MMAP -LANG=OTHER -NEG_SCORES=0 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sta:*
28: em_un:*

29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_hum:*
32: em_hgt_hum:*
33: em_hgt_hum:*
34: em_hgt_hum:*
35: em_hgt_hum:*
36: em_hgt_hum:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgt_hum:*
40: em_hgt_hum:*
41: em_hgt_hum:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	350	100.0	1053	6 AX119075	AX119075 Sequence
2	350	100.0	1387	9 AF060513	AF060513 Homo sapi
3	350	100.0	1425	9 AF283538	AF283538 Homo sapi
4	350	100.0	1440	9 BC007654	BC007654 Homo sapi
5	350	100.0	1574	6 BD158389	BD158389 Homo sapi
6	350	100.0	1574	9 AK023378	AK023378 Homo sapi
7	350	100.0	1606	9 AB038949	AB038949 Homo sapi
8	332	94.9	1353	9 HSM05532	HSM05532 Homo sapi
9	311	88.9	1579	9 AF182416	AF182416 Homo sapi
10	122	34.9	796	6 BD149184	BD149184 Primer fo
11	119	34.0	170586	2 AC037455	AC037455 Homo sapi
12	119	34.0	190508	2 AC005037	AC005037 Homo sapi
13	119	34.0	198250	2 AC093681	AC093681 Homo sapi
14	70	20.0	249	6 BD049005	BD049005 Sequence
15	53	15.1	231600	2 AC130779	AC130779 Rattus no
16	51	14.6	1836	10 AF284439	AF284439 Mus muscu
17	51	14.6	149819	2 AC121091	AC121091 Mus muscu
18	51	14.6	234976	2 AC118698	AC118698 Mus muscu
19	46	13.1	198250	2 AC093681	AC093681 Homo sapi
20	24	6.9	231600	2 AC130779	AC130779 Rattus no
21	24	6.9	281017	2 AC123462	AC123462 Rattus no
22	17	4.9	1328	3 AK114307	AK114307 Clona int
23	16	4.6	55374	9 AL645474	AL645474 Human DNA
24	14	4.0	155127	2 AL929495	AL929495 Danio rer
25	10	2.9	339	6 BD040994	BD040994 Sequence
26	10	2.9	1083	6 AX143427	AX143427 Sequence
27	10	2.9	1098	6 AX617526	AX617526 Sequence
28	10	2.9	3041	6 AF270328	AF270328 Staphyloc
29	10	2.9	3041	6 AX145646	AX145646 Sequence
30	10	2.9	3138	1 AF269913	AF269913 Staphyloc
31	10	2.9	3188	6 AX145231	AX145231 Sequence
32	10	2.9	3441	1 AF269767	AF269767 Staphyloc
33	10	2.9	3441	6 AX145085	AX145085 Sequence
34	10	2.9	171193	10 AC123040	AC123040 Mus muscu
35	10	2.9	300028	1 AE016748	AE016748 Staphyloc
36	10	2.9	301550	1 AP003134	AP003134 Staphyloc
37	10	2.9	333750	1 AP004827	AP004827 Staphyloc
38	10	2.9	346900	1 AP003362	AP003362 Staphyloc
39	9	2.6	414	11 DM3H2S	DM3H2S
40	9	2.6	901	8 SCUS2042	SCUS2042 Saccharomyc
41	9	2.6	1187	8 SCY61221C	SCY61221C
42	9	2.6	1248	6 AX608648	AX608648 Sequence
43	9	2.6	4946	5 TR0544607	TR0544607 Takifugu
44	9	2.6	10979	1 AE009917	AE009917 Pyrococcus
45	9	2.6	13617	1 AE010179	AE010179 Pyrococcus

RESULT 1

ALIGNMENTS

AX119075
LOCUS AX119075 1053 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 239 from Patent WO0129221.
ACCESSION AX119075
VERSION AX119075.1 GI:14036029
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Conklin, D.C. and Yee, D.P.
Proteins and polynucleotides encoding them
Patent: WO 0129221-A 239 26-Apr-2001;
ZymoGenetics, Inc. (US)
FEATURES
source location/Qualifiers
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..1053
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC38611.1"
/translation="MDIKALSSINDRASISFASWNVGLVPEPPHYNTLEFTH
DLTEVMEVYQKRAADLLSYHPFIFPMKRTITWNTWERTVIALERHVTSPHTA
YDAAPQGVNMLAKGAGCTSRPIHPSKAPVPEGNHRENNVNTODDVKVSAVK
GIDGVSYSFASRTGNEQRIINLNCQKALMVYDFSRKQLOKTEILSLKPL
LHTMGRCITLDESVSLATMIDIRKHLKISHIRALGVGRTLSQVYVVALAGSGS
SVLQGVADYITLGEVMSHDTLDAASQINVIICEHSTENGFLSDLDMDLSHLENK
INITLSEDRPFLVYV"
BASE COUNT 279 a 253 c 247 g 274 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1053
Score: 350.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-745-506-37 (1-350) x AX119075 (1-1053)
QY 1 MetaspLeuLysAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
DB 1 ATGATTTGAAGGCTCTCTTCTCTTGAATGATGATCCCTCTCTCTCTGAG 60
QY 21 SerTrpAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
DB 61 AGTTGGAGCAATGTTGGATTACTGTGAGACCAACCCACACATCTGTAATACACATC 120
QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLysAlaAsp 60
DB 121 TTCTCGACCAATGAGCTCTGAGAGAGATGAGAGGCTCTCAAAAAGAGCAGAC 180
QY 61 LeuLeuSerTyrHisProProLeuPheArgProMetLysArgIleThrTrpAsnThr 80
DB 181 CTCATCTCTCTCTACATCCGCGCTATCTTCGACCAATGAGAGCATATACTGGAACACA 240
QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
DB 241 TGAAGAGAGCGCTGATGATCGGAGCTGTGAGAACAGAGTGTATCTACTCTCTCAT 300
QY 101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuAla 120
DB 301 ACAGGCTATGATGCTGCGCCGAGGCGTCAACAACATGGTGGCTAAAGGCGCTGAGCT 360
QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGluLysAsnHis 140
DB 361 TGTACCTCGAGGCCATACATCTCTTCCAAAGCTCCCACTACCTACAGAGGAAACCCAC 420

QY 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
DB 421 CGAGTAAATTCACAGCTTACTACACCCAGACCTGAGCAAAAGTCATCTCGCATAA 480
QY 161 GlyLeuAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
DB 481 GGAATTGACGGTGTCTTCTGTCATCTTTTCTCTCTAGACTGGTATGAGGAACAACA 540
QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
DB 541 CGGATTAATCTGGAATTGTACTACAGAGGCTTGTATGACAGGCTGATATTTCTTCCCGG 600
QY 201 AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
DB 601 AACAAACAACTTTATCAGAAACGGAATTTCTGACCTGAGAAAGCCCTTGCTTTCAT 660
QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetLeuAsp 240
DB 661 ACTGGAATGGAGCGCTTATGACACACTGATGAAATCTGCTCCCGGACACCATGATGAT 720
QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
DB 721 CGAATTAACAAACACACCTAAACATCTCATATTCGCTTACGCCCTTGGGCTGGGAGAACC 780
QY 261 LeuGluSerGlnValLysValValAlaLeuCysAlaGlySerGlySerValLeuGln 280
DB 781 TTAGAGTCTCAAGTCAAAGTCAGTGGCCCTGTGCTGCTGTGGAGACAGGCTCTGACG 840
QY 281 GlyValGluAlaAspLeuTyrLeuThrGlyLysMetSerHisAspThrLeuAspAla 300
DB 841 GGTGTGAGGCTGACCTTACTCTCAGAGTGAATGCTCCATCATGATACTTGGATGCT 900
QY 301 AlaSerGlnGlyIleAsnValIleLeuCysGlnHisSerAsnThrGluArgGlyPheLeu 320
DB 901 GCTTCCCAAGAAATAATGATCATCTCTGTGAACACACAACTGAACGAGGCTTCTT 960
QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleLeuSer 340
DB 961 TCTGACCTTCAGATATCTCTGATTTCTCACTTGGAGATTAAGTAATATATCTCATCA 1020
QY 341 GluThrAspArgAspProLeuGlnValVal 350
DB 1021 GAGACTGACAGGACCTCTTCAGGTGCTA 1050
RESULT 2
AF060513 1387 bp mRNA linear PRI 02-JAN-2001
LOCUS AF060513
DEFINITION Homo sapiens clone 016d06 My018 protein mRNA, complete cds.
ACCESSION AF060513
VERSION AF060513.1 GI:12001975
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1387)
Mao, Y.M., Xie, Y., Huang, X.Y., Ying, K. and Dai, J.L.
Direct Submission
Submitted (20-Apr-1998) Institute of Genetics, School of Life
Science, Fudan University, 220 Handan Rd., Shanghai 200433,
P.R.China
FEATURES
source location/Qualifiers
1..1387
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="016d06"
/tissue_type="brain"
/dev_stage="fetus"
82..1134
/codon_start=1
/product="My018 protein"
/protein_id="A043131.1"

/db_xref="GI:12001976"
 /translation="MDLKLALSLNDFASLSPAESMDVNGLVLEPSPPTVNTFLTN
 DLTEEMEEVLQKADLILSYHPIERPKRTMTWMTKRLVIRALENVGIIYSPITA
 YDAPOGVNMLAKIGIAGCTSPRIHPSKAPNPTBESNHEVENVNTQDLDVMSAVK
 GIDGVSVTSFASRTGNEQTRINLNCQALQVDFLSRNOQYOKETILSLERKL
 LHTGVRCTLDESIVLATMIRKRLHSIRLALGVRILGSLGVKVAALACGSGS
 SVLQGVADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGEFLSDLRMLDSLENK
 INITSEMDRDLQVY"
 BASE COUNT 378 a 323 c 309 g 377 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 1387
 Score: 350.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x AF060513 (1-1387)

QY 1 MetAplLeuValAlaLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 82 ATGGATTGGAAGGCTCTCTCTCTCTGGAATGACTTGACCTCCCTCTGTTGCTGAG 141
 QY 21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
 DB 142 AGTTGGACAAATGTTGATTACTGCGGAGAACCAAGCCCAACATCTGTAATACACTC 201
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluLeuGluValAsp 60
 DB 202 TTCCCTGACCAATGACCTGACTGAGCAAGATGAGAGAGCTGCTGCAAAAGAGCAGAC 261
 QY 61 LeuIleLeuSerThrHisProProIlePheArgProMetLeuArgIleThrPAsnThr 80
 DB 262 CTCATCTCTCTCTACCATCCGCTATCTCCGACCCATGAGGCGCTAACCTGGAGACACA 321
 QY 81 TPlySGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleThrSerProHis 100
 DB 322 TCGAAGAGAGCGCGCTGGTATCCGGGCTCGGAGAAAGAGTGGTATCTACTCTCAT 381
 QY 101 ThrAlaIleThrAspAlaAlaProGlnGlyValAsnAsnTPLeuAlaGlyLeuGlyAla 120
 DB 382 ACAGCTATGATGATGCGCCGCCAGGCGCTCAACACTGCTGCTAAAGGCTTGGAGCT 441
 QY 121 CysThrSerArgProIleHisProSerLeuAlaProAsnTPProThrGluGluAsnHis 140
 DB 442 TGTACCTCCAGGCCCTACATCTCCAAAGCTCCCACTACCTACCAAGGGAACACAC 501
 QY 141 ArgValGluPheAsnValAsnTPThrGlnAspLeuAspLeuValMetSerAlaValHis 160
 DB 502 CGAGTAGAATTCACAGCTTAACACCCCAAGACCTGGCAAGATCATGTCGAGGAAA 561
 QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluGlnThr 180
 DB 562 GGAATTTGAGGCTGTTCTGCTACTCTTTCTGCTAGACCTGGTATGAGGAACAAACA 621
 QY 181 ArgIleAsnLeuAsnTPThrGlnLeuValAlaLeuMetGluValAlaAspPheLeuSerArg 200
 DB 622 CGGATTATCTGAATTTGTTACTCAGAAAGCTTTGATGAGGTGATGATTTCTTCCCG 681
 QY 201 AsnLeuGlnLeuTPGlnLeuTPThrGluIleLeuSerLeuGluLeuProLeuLeuLeuHis 220
 DB 682 AACCAAAACAATTATCAGAAAGGAAATTCGTCTACTGAGAGAGCTTTGGCTTACAT 741
 QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
 DB 742 ACTGGAAATGGACGGTTATGACACACTGATGATGATCTGCTCCCTGGCAACATGATGAT 801
 QY 241 ArgIleLeuArgHisLeuLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
 DB 802 CGAATTAAGACACCTTAACATCTCATATTCGTTAGCCCTTGGGGGGGAGAACCC 861

QY 261 LeuGluSerGlnValIleValAlaLeuCysAlaGlySerGlySerValLeuGln 280
 DB 862 TTGAGCTTCACATCAAGTCGTGGCCCTTGCTGCTGCTTCTGGAGACAGCTTCTGCAG 921
 QY 281 GlyValGluAlaAspLeuTPLeuThrGlyGluMetSerHisIleAspThrLeuAspAla 300
 DB 922 GGTGTTGAGGCTCACTTACCTACACAGGAGATGTCACATCATGATATTATCTTATCA 981
 QY 301 AlaSerGlnGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgIleLeu 320
 DB 982 GCTTCCCAAGGAATTAATGTCATCTCTGGAACACAGCAACACTGAAACGAGCTTCTT 1041
 QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLeuIleLeuSer 340
 DB 1042 TCTGACCTTCGAAATATGCTGATCTCTCACTTGGAGATAGATTAATATTATCTTATCA 1101
 QY 341 GluThrAspArgAspProLeuGlnValAla 350
 DB 1102 GAGACTGACAGGACCCCTTTCAGTGTGTA 1131

RESULT 3
 AF283538
 LOCUS AF283538 1425 bp mRNA linear PRI 16-JAN-2001
 DEFINITION Homo sapiens NIF3L1 protein mRNA, complete cds.
 ACCESSION AF283538
 VERSION AF283538.1 GI:12006402
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Tascou S., Burfeind P., and Engel W.
 TITLE 1 (bases 1 to 1425)
 Burfeind P., Uedelhoven J., Dijkens C., Nayerina K., Engel W. and
 Isolation and characterization of a novel human gene, NIF3L1, and
 its mouse ortholog, Nif3l1, highly conserved from bacteria to
 mammals

JOURNAL
 MEDLINE
 PUBMED 20573864
 11124544
 REFERENCE
 2 (bases 1 to 1425)
 Tascou S., Burfeind P., and Engel W.
 TITLE Direct Submission
 JOURNAL Submitted (29-JUN-2000) Institute for Human Genetics, University of
 Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany

FEATURES

source
 1..1425
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 47..1180
 /codon_start=1
 /product="NIF3L1 protein"
 /protein_id="AA044846.1"
 /db_xref="GI:12006403"

CDS

/translation="MSSCVRPVPTVPRFVDSLICNSRSFMDKLISLNDFAIS
 FASMDNVGLVPPSPPHVNTFLNDLTERMEVLOKKADLISYHPIERPKR
 ITVNTKERLYTALENRGVTSPTATYDAAPGVNMLAKIGIAGCTSPRIHPSKAPN
 YPEGHNHREFFNVNNTQDLDKVASAVKIDGVSVTSFASRTGNEQTRINLNCQAL
 MOVNDFLSRNKOLYOKETILSLERPLILHTGMRGLCTLDESIVLATMIRKRLHS
 HIRLALGVRILGSLGVKVAALACGSGSVLQGVADLYLTGEMSHHDTLDAASQGINV
 ILCEHSNTERGEFLSDLRMLDSLEKNINIISETDRDLQVY"

BASE COUNT 385 a 331 c 314 g 395 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 1425
 Score: 350.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

```

US-09-745-506-37 (1-350) x AF283538 (1-1425)
QY      1 MetaspLeuylsAlaleuSerSerleuAsnAspPheAlaSerleuSerPheAlaGlu 20
      128 ATGATATTGAAGGCTCTCTTCTCTTCTTGAATGACATTCACACCCCTGCTTGGCTGAG 187
QY      21 SerTPAspAsnValGlyLeuLeuValGluProSerProHisThrValAsnThrLeu 40
      188 AGTTGGACAAATGTTGATTACTGCTGGAAACCAAGCCACACATCTGTAATACACTC 247
QY      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp 60
      248 TTCTTGACCAATGCTGACTGAGAGATGATGAGAGAGTCTGCAAAAAGAGCGAC 307
QY      61 LeuLeuSerTyRHisProPheArgPheArgPheMetLysArgLierThrPAsnThr 80
      308 CTCATTCTCTCCATACCATCGCCATCTCCGACCATGAAAGCCATACCTGGAACACA 367
QY      81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyRSerProHis 100
      368 TGGAGAGAGCGCGCTGGATCCGGGCTGAGAGAACAGAGTCGTTACTCTCTCCAT 427
QY      101 ThrAlaArgAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGlyAla 120
      428 ACAGCTATGATGCTGCGCCCGAGGCGCTCAACACTGTTGGCTTAAAGGCTTGGAGCT 487
QY      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyRProThrGluGlyAsnHis 140
      488 TGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCAACTACCTACAGAGGAAACAC 547
QY      141 ArgValGluPheAsnValAsnTyRThrGlnAspLeuAspLysValMetSerAlaValLys 160
      548 CGAGTACAATTCACAGTTAACTACACCAAGACCTGAGCAAAAGTCATGCTGCAGTGA 607
QY      161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
      608 GGAATTACGCGTGTCTGCTACTCTCTCTCTGAGACTGATGAGAAACAAACA 667
QY      181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
      668 CGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
QY      201 AsnLysGlnLeuTyRArgLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
      728 AACCAACAACTTATCTGAGAACGAGAAATCTGTCACGAGAGAGCCCTTCTCTACAT 787
QY      221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
      788 ACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
QY      241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValArgThr 260
      848 CGAATTAACAAACACACCTTAACCTATCTGATGATGATGATGATGATGATGATGAT 907
QY      261 LeuGluSerGlnValLysValValAlaLeuCysAlaGlySerGlySerSerValLeuGln 280
      908 TTAAAGTCTCAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
QY      281 GlyValGluAlaAspLeuTyRLeuThrGlyGluMetSerHisAspThrLeuAspAla 300
      968 GGTGTGGAGCTGACCTTACCTACAGAGTATGATGATGATGATGATGATGATGATGAT 1027
QY      301 AlaSerGlnGlyIleAsnValIleLeuGlyGlnHisSerAsnThrGluArgGlyPheLeu 320
      1028 GCTTCCCAAGAAATAATGATCTCTGTCGACACACACACACACACACACACACACACAC 1087
QY      321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340
      1088 TCTGACCTTCAGAGATGCTGATCTCATTGAGAAATGATTAATATATATCTATCA 1147
QY      341 GluThrAspArgAspProLeuGlnValVal 350
      |||||||

```

```

Db      1148 GAGACTGACAGGAGACCTCTTCAGTGTGTA 1177
RESULT 4
LOCUS   BC007654
DEFINITION
Homo sapiens, similar to NIF3 (Nrg1 interacting factor 3, 5.pombe homolog)-like 1, clone MGC:2595 IMAGE:3346099, mRNA, complete cds.
ACCESSION
BC007654.1 GI:14043316
VERSION
BC007654
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
Organism
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1440)
Strausberg, R.
Direct Submission
Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCFP/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Boultad, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J., Tlonsgon, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov
Series: IRAL Plate: 5 Row: 1 Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10197631.
FEATURES
source
Location/Qualifiers
1..1440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:2595 IMAGE:3346099"
/tissue_type="Skin, melanotic melanoma."
/clone_11b="NIF3_MGC_20"
/lab_host="DH10B-R"
/note="vector: pOTB7"
37..1170
/codon_start=1
/product="Similar to NIF3 (Nrg1 interacting factor 3, 5.pombe homolog)-like 1."
/protein_id="AAH07654.1"
/db_xref="GI:14043317"
/translacion="MISCVRPPTVTFVFDLSLNCSSRSFMDLKALISLNDASLTS
FAESMDNGLVLESPRPVTVTFVFDLSLNCSSRSFMDLKALISLNDASLTS
ITWNTWKRLEVTALERNVGYISPTATDAPGCVNNMLARGCLASCSRPVLPKAPN
YPTGNNHVEVNVNTODLDVMSAVKIDGVSVTSFARIGNEBOTRINTNCOKAL
MOVVDLSRNLOKTEILSLERPLTLHTMGRLCTDESVSLATMDLRKRLKLS
HIRLALGRTLESQVYVALCAGSSVLDGVEDLVLTGEMSHHDTLDAASOGINV
ILCESNTERGFISDLRMDLSLENKNIILISFDRDPLGVV"
BASE COUNT
407 a 329 c 313 g 391 t
ORIGIN
Alignment Scores: 0 Length: 1440
Pred. No.:

```

Score: 350.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x BC007654 (1-1440)

QY 1 MetAspLeuValAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 118 ATGGATTGAAAGGCTCTCTCTCTCTGAAATGACTTGACCTCTCTCTGCTGAG 177
 QY 21 SerTPAspAsnValAlaGlyLeuLeuValGluProSerProPheHisThrValAsnThrLeu 40
 DB 178 AGTTGGACAATGTTGGATTACTGGTGGAAACAGCCCAACATATCTGTAATACACATC 237
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluGluValAlaAsp 60
 DB 238 TTCTGACCAATGACCTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
 QY 61 LeuIleLeuSerTyHisProPheArgProMetLeuArgIleThrTrpAsnThr 80
 DB 298 CTCATTCT 357
 QY 81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValAlaGlyIleTyrSerProHis 100
 DB 358 TGGAG 417
 QY 101 ThrAlaIleTyrAspAlaAlaProGluGluValAsnAsnTrpLeuAlaLysGlyLeuGluAla 120
 DB 418 ACAGCTTGTATGATGCTGCGCCGAGGCGCTCAACCACTGCTGCTAAAGGCTTGGAGCT 477
 QY 121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGluGluHis 140
 DB 478 TGTACTCTCCAGGCT 537
 QY 141 ArgValGluPheAsnValAsnTyrThrGluAspLeuAspLysValMetSerAlaValLys 160
 DB 538 CGAGTAGAATTCAGCTTACCTACACCCCAAGACCTGAGCAAAAGTCTGCTGAGTGA 597
 QY 161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluGlnThr 180
 DB 598 GGAATGACGGGTGCTGCT 657
 QY 181 ArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGluValValAspPheLeuSerArg 200
 DB 658 CGGATTAACTGAATTCATCTCAGAAAGCTTTGATGACAGCTGGTAGATTCTTCCCGG 717
 QY 201 AsnLysGluLeuTyrGlnLysThrGluIleLeuSerLeuGluLysPheLeuLeuHis 220
 DB 718 AACCAACAACCTTATCAGAAAGCAAAATTCGTCTGCTGAGAGAGAGAGAGAGAGAGAG 777
 QY 221 ThrGlyMetGlyArgLeuCysThrLeuAspGluSerValSerLeuAlaThrMetLeuAsp 240
 DB 778 ACTGGAATGAGAGGCTTATGACACTGATGATGATGATGATGATGATGATGATGATGAT 837
 QY 241 ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGluValAlaGlyArgThr 260
 DB 838 CGAATAAAAAGACACTAAATATCTATCTATCTGCTTATGCTTGGGGGAGAGAGAGAG 897
 QY 261 LeuGluSerGluValLysValAlaLeuCysAlaGlySerGlySerSerValLeuGlu 280
 DB 898 TTAGAGTCTCAGTCAAAAGTCTGGGCGCTGTGCTGGTGGGAGAGAGAGAGAGAGAGAG 957
 QY 281 GlyValIleGluAlaAspLeuTyrIleThrGlyGluMetSerHisHisAspThrLeuAspAla 300
 DB 958 GGTGTGGAGGCTGACCTTATCTCAGAGTGAATGATGATGATGATGATGATGATGATGAT 1017
 QY 301 AlaSerGluGlyIleAsnValIleLeuCysGluHisSerAsnThrGluArgGlyPheLeu 320
 DB 1018 GCTTCCCAAGAAATTAATGCTATCTCTGTGAAACAGCAACATGAAAGAGAGCTTCTT 1077
 QY 321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340

DB 1078 TCTGACCTCGAGATATGCTGATTTCTCACTTGGAGAAATATATATATATATATATAT 1137
 QY 341 GluThrAspArgAspProLeuGluValAla 350
 DB 1138 GAGACTGACAGGAGGACCTCTCTCAGGTGGA 1167

RESULT 5

LOCUS BD158389 1574 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD158389
 VERSION BD158389.1 GI:27864147
 KEYWORDS JP 2002191363-A/13232.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Ota,T., Isozaki,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1574)

TITLE
 JOURNAL
 Helix RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/13232
 PD 09-JUL-2002

COMMENT
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOZAKI, TETSUO NISHIKAWA, KOJI HAYASHI, KIORU
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES
 source FT CDS Location/Qualifiers
 1..1574
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 (271)..(1320).

BASE COUNT 420 a 361 c 372 g 421 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 1574
 Score: 350.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-745-506-37 (1-350) x BD158389 (1-1574)

QY 1 MetAspLeuValAlaLeuLeuSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
 DB 271 ATGGATTGAAAGGCTCTCTCTCTCTGAAATGACTTGACCTCTCTCTGCTGAG 330
 QY 21 SerTPAspAsnValAlaGlyLeuLeuValGluProSerProPheHisThrValAsnThrLeu 40
 DB 331 AGTTGGACAATGTTGGATTACTGGTGGAAACAGCCCAACATATCTGTAATACACATC 390
 QY 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGluGluValAlaAsp 60
 DB 391 TTCTGACCAATGACCTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
 QY 61 LeuIleLeuSerTyHisProPheArgProMetLeuArgIleThrTrpAsnThr 80
 DB 451 CTCATTCT 510

```

OY      81  TTrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
        |||
        |||
        |||
Db      511  TGGAGAGAGCGCGCTGGATCCGGGCTCTGGAGAACAGAGCGGTATCTACTCTCCAT 570
OY      101  ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaIlyGlyLeuGlyAla 120
        |||
        |||
        |||
Db      571  ACAGCCATATATGCTGGCGCCCGAGGGGCTCAACAACGCTGGTCTAAAGGCTTGGAGCT 630
OY      121  CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGlnGlyAlaHis 140
        |||
        |||
        |||
Db      631  TGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCAACTACCTTACAGAGGGAACAC 690
OY      141  ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
        |||
        |||
        |||
Db      691  CGAGTAGAATTCACAGCTTAACCTACACCAAGACCTGGACAAGATCATGCTGCAGTGAAA 750
OY      161  GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
        |||
        |||
        |||
Db      751  GGAATTCAGCGGTGTTGTGCACCTCTTTCTGCTGGAGCTGGTAATGAGAACCAACA 810
OY      181  ArgIleAsnLeuAsnGlyThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
        |||
        |||
        |||
Db      811  CGGATTATATCGATTCTACTACAGAGGCTTGTGATGACAGGTGATGATTTCTTCCCG 870
OY      201  AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGlnLysProLeuLeuHis 220
        |||
        |||
        |||
Db      871  AACCAACACCTTATATCAGAAAGACGAAATCTGTACATGGAGAACCTTGTCTCTACAT 930
OY      221  ThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetLeuAsp 240
        |||
        |||
        |||
Db      931  ACTGGAAATGGAGCGGTATGACACCTGAGTAATCTGTCTCTCCGACCAACATGATGAT 990
OY      241  ArgIleLysArgHisLeuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
        |||
        |||
        |||
Db      991  CGAATAAAGAGACACCTAAACATCATATTCGCTTACGCCCTTGGGGGTGGAGAAC 1050
OY      261  LeuGluSerGlnValLysValValAlaLeuGlyAlaGlySerGlySerSerValLeuGln 280
        |||
        |||
        |||
Db      1051  TTAAAGCTCTCAAGTCAAGACGCTGGCCCTGTGTCTGTCTGGAGACGCTTCTGCAG 1110
OY      281  GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisAspThrLeuAspAla 300
        |||
        |||
        |||
Db      1111  GGTGTGAGGCTGACCTTTTACCTCAGAGTGAATGTCCTCCATCATGATCTTGGATGCT 1170
OY      301  AlaSerGlnGlyIleAsnValIleLeuGlyGlnHisSerAsnThrGluArgGlyPheLeu 320
        |||
        |||
        |||
Db      1171  GCTTCCCAAGAAATAATGTCATCTCTGTGACACACGACACACTGAACGAGGCTTCTT 1230
OY      321  SerAspLeuArgAspMetLeuAspSerHisLeuGlnLysLysIleAsnIleLeuSer 340
        |||
        |||
        |||
Db      1231  TCTGACCTTCGAGATATGCTGGATTCACCTTGGAGATATAGATTAATATCTATCA 1290
OY      341  GluThrAspArgAspProLeuGlnValVal 350
        |||
        |||
        |||
Db      1291  GAGACTGACAGGAGACCTCTTCAGGTGTA 1320

```

```

RESULT 6
AK023378 1574 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ1316 f1s, clone OVARC1001555, weakly similar
DEFINITION to NGFI-INTERACTING FACTOR 3.
ACCESSION AK023378
VERSION AK023378.1 GI:10435289
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Mugatauma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,

```

```

OYO, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,
Nakamura, Y., Nagahara, K., Masuno, Y., Minomiyama, K. and Iwayanagi, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1574)
REFERENCE
AUTHORS Isogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory: 1532-3 Yano, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..1574
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OVARC1001555"
/tissue_type="ovary, tumor tissue"
/clone_id="OVARC1"
/feature="cloning vector: pME18SFL3"
271..1323
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14551.1"
/db_xref="GI:10435290"
/translating="MDKALISLNDPASFASBMDNVLVEPSDPHTVNTLELN
DLREVMEEVLOKRAADILSYHPFIPFMRKTRTWKERYVIRALNRVGIYSPHRA
YDAPDOGVNMMALGACGCTSRPSPKAPVPTGHNHREYFNVDIDDKVSAVK
GIDGVSPTSAPRNGEORFNRINCPKOKALMOYVDNLSRKLQYKTEILSEKPL
LHFMGRICITDESYSLATMADRKRHLKSHITLALGCTGRTESQKYVALCGSS
SVLQGVADLITLTEMESHDTLDASGGINVITLDEHNTERTGFLSDRLMDLSLEMK
INITLSDRLPLQV"
BASE COUNT 420 a 361 c 372 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1574
Score: 350.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-745-506-37 (1-350) x AK023378 (1-1574)
OY      1  MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
        |||
        |||
        |||
Db      271  ATGCAATTGGAAGCGCTCTCTCTTCTTGAATGACTTGTGATCCCTCTGTTGGCTAG 330
OY      21  SerTrpAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
        |||
        |||
        |||
Db      331  AGTTGGGCAATATGTTGATCTGCTGGACCAAGCCCAACACATACTGTAATACACTC 390
OY      41  PheLeuThrAsnAspLeuThrGlnGluValMetGluGluValLeuGlnLysLysAlaAsp 60
        |||
        |||
        |||
Db      391  TTCTGACCAATGACCTGAGAGAGAGTGAATGAGGAGGCTGCCAAAGAGGACAGAC 450
OY      61  LeuIleLeuSerTyrHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80
        |||
        |||
        |||
Db      451  CTGATTTCTCTCTACATCCCGCTATCTTCCGACCAAGAGGCGATACCTGAGACACA 510
OY      81  TTrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
        |||
        |||
        |||
Db      511  TGGAGAGAGCGCGCTGGATCCGGGCTCTGGAGAACAGAGCGGTATCTACTCTCCAT 570
OY      101  ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaIlyGlyLeuGlyAla 120

```

```

Db      571 ACAGCCTATGATGCTCGCCCGGAGGCGTCACAACTGTTGGCTAAAGGCTTGAGCT 630
Qy      121 CysThrSerArgProIleHisProSerIysAlaProAsnTyrProThrGluGlyAsnHis 140
Db      631 TGACTCTCCAGGCGCCACATCCCTTCCAAAGCTCCCAACTACCTCAGAGCGGAAACAC 690
Qy      141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspIysValMetSerAlaValLys 160
Db      691 CAGATGAAATTCACCTTAACCTACACCAAGACCTGGACAAAGTCACTGCTGCAGTGA 750
Qy      161 GlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGluThr 180
Db      751 GGAAATGAGCGGTCTTCTGCTCACTCTTTTCTGCTAGGACTGTAATGAGAAACA 810
Qy      181 ArgIleAsnLeuAsnGlyThrGlnLysAlaLeuMetGlnValValAspPheLeuSerArg 200
Db      811 CGGATTAACTCGAATTTGTAAGGCTTGTATGATCAGGTGATGATTTCTTCCCGG 870
Qy      201 AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluIysProLeuLeuHis 220
Db      871 AACAAACAACTTATATCAGAAAGCGAAATTCCTCTCAGTGAAGAACCTTTGCTTACAT 930
Qy      221 ThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetIleAsp 240
Db      931 ACTGGAATGGAGCGGTATGACACACAGATGATGATCTGCTCCCTGGCAACATGATGAT 990
Qy      241 ArgIleLysArgHisIleuLysLeuSerHisIleArgLeuAlaLeuGlyValGlyArgThr 260
Db      991 CGAATTAAGAAAGACCTTAAACATCTCATATTCCCTTACCCCTGGGGGGGAGAAACC 1050
Qy      261 LeuGluSerGlnValLysValValAlaLeuGlyAlaGlySerGlySerValLeuGln 280
Db      1051 TTAGACTCTCAAGTCAAAGCGCGGCGCTGTGCTGTGCTGTGGAGACCGTTCTGCAG 1110
Qy      281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisIleAspThrLeuAspAla 300
Db      1111 GGCTGTGAGGCTGACCTTACCTCACAGGAGATGCCATCATGATCTTGGATGCT 1170
Qy      301 AlaSerGlnGlyIleAsnValIleLeuGlyGluHisSerAsnThrGluArgGlyPheLeu 320
Db      1171 GCTTCCCAAGGAATTAATGTCATCTGTGAACACAGCAACCTGAAACGAGCTTCTT 1230
Qy      321 SerAspLeuArgAspMetLeuAspSerHisLeuGluAsnLysIleAsnIleIleLeuSer 340
Db      1231 TCGACCTTCGAGATATGCTGCTCTCCTTCTGAGATGATGATTAATTAATTCCTATCA 1290
Qy      341 GluThrAspArgAspProLeuGlnValVal 350
Db      1291 GAGACTGACAGGAGCCCTTCCTGAGTGTA 1320

RESULT 7
AB038949 1606 bp mRNA linear PRI 23-FEB-2001
LOCUS AB038949 candidate 1, complete cds.
DEFINITION Homo sapiens ALS2CRL mRNA for amyotrophic lateral sclerosis 2.
ACCESSION AB038949
VERSION AB038949.1 GI:12862477
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE Hadano,S., Yanagisawa,Y., Skaug,J., Fichter,K., Nasir,J.,
AUTHORS Marlandale,D., Koop,B.F., Scherer,S.W., Nicholson,D.W.,
Rouleau,G.A., Ikeda,J.-E. and Hayden,M.R.
ROUTING cloning and characterization of three novel genes, ALS2CRL,
ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis
(ALS2) critical region at chromosome 2q33-q34: candidate genes for
ALS2
JOURNAL Genomics 71 (2), 200-213 (2001)
MEDLINE 21100893

```

```

PUBMED 11161814
REFERENCE 2 (bases 1 to 1606)
AUTHORS Hadano,S., Ikeda,J. and Hayden,M.R.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Shunji Hadano, Tokai University, The
Institute of Medical Sciences, Bohseidai, Isehara, Kanagawa
259-1193, Japan (E-mail:shunji@eng.med.u-tokai.ac.jp,
Tel.:+81-463-91-5095, Fax:+81-463-91-4993)
FEATURES
source
1.1606
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q33"
/note="IMAGE:32584-Unigene ID: Hs.21943"
1.1606
/gene="ALS2CRL"
286..1338
/gene="ALS2CRL"
/note="candidate 1"
/codon_start=1
/product="amyotrophic lateral sclerosis 2"
/protein_id="BAB32499.1"
/db_xref="GI:12862478"
/translation="MDKLALSLNDPASLSPAESMDNVGLVPSPPHTVNTLFLTN
DLPEYMEVLOKKADLLSYHPFRPMKRTVMWKERYVIRALENRYGYSPTA
YDAAPGVNMLLAKGLACTSRPIHRSKAPNTTEGHNHREYVNTODLDKYSAYK
GIDGVSVTSFASRTGNEEDRIINLCTOKALMVDPFLVSKKOLYOTETLSLEKPL
LHTGMORLCTLDSESVLATMIDIRIKRLKLSHLRLAGVRLTESOVKVALCAGGS
SVLGVEADLYLTGEMSHDPTLDAASQGINVILCEHSNTERGFLSLRDLMDLSLENK
INILSEDDDPQVY"
1566..1571
polya_signal
BASE COUNT 438 a 365 c 375 g 428 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 1606
Score: 350.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-745-506-37 (1-350) x AB038949 (1-1606)
Qy      1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
Db      286 ATGGAATTGAAGGCTCTCTTCTCTCTTGAATGACTTTCGATCCCTCGTTGCTGAG 345
Qy      21 SerTrpAspAsnValGlyLeuLeuValGluProSerProThrValAsnThrLeu 40
Db      346 AGTTGGGACAATGTTGGATTACTGTGGAACCAAGCCACACATACCTGTAATACACTC 405
Qy      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysLysAlaAsp 60
Db      406 TTCTTGACCAATACCTGACTGAGGAAGTATGAGAGGAGTCTGCAAAAGAACGCAGAC 465
Qy      61 LeuIleLeuSerTyrHisProPheArgProMetLysArgIleThrTrpAsnThr 80
Db      466 CTCATCTCTCCCTACCACTCCGCTATCTTCGACCCATGAGCGCATACCTGGACACA 525
Qy      81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
Db      526 TGAAGAGAGCGCTGGATGATCGGAGCTGTGAACACAGAGATCCGATCTACTCTCCAT 585
Qy      101 ThrAlaTyrTrpAlaIleArgGlnGlyValAsnAspTrpLeuAlaLysGlyLeuGlyAla 120
Db      586 ACAGCCTATGATGCTCGCCCGGAGGCGTCAACAACTGCTGCTAAAGGCTTGAGCT 645
Qy      121 CysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGlyAsnHis 140

```


Db 646 TGTACTCCAGCCCAATACATCTTCCAAAGCTCCCACTACCTACAGAGGGAACCAAC 705
 Qy 141 ArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaValLys 160
 Db 706 CGAGTACAAATTCAGATTAACCTACACCAAGACCTGGACAAGACATGCTGCAGTGAAA 765
 Qy 161 GylTLeaspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGlnThr 180
 Db 766 GGAATTCACGGTGTCTGTCTACTCTTTCTCTCTAGACTGTGTAATGAGAACCAACA 825
 Qy 181 ArgTLeasLeuAsnGlyThrGlnLysAlaLeuMetGlnValAlaAspPheLeuSerArg 200
 Db 826 CGGATTAATCTGAATTTGATCTACAGAGCTTTGATGACAGGAGTATTTCTTCCCG 885
 Qy 201 AsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeuHis 220
 Db 886 AACCAACAACTTATTCAGAAAGACGAAATTCGTACAGAGAGACCTTCTCTCAT 945
 Qy 221 ThrGlyMetGlyAlaGlyLeuGlyThrLeuAspGluSerValSerLeuAlaThrMetLys 240
 Db 946 ACTGGAATGGACGGTATGACACATGATGAAATCTGCTCCCGCAACCATGATTCAT 1005
 Qy 241 ArgTLeasArgHisLeuLysLeuSerHisTleArgLeuAlaLeuGlyValArgThr 260
 Db 1006 CGAATTAACAAAGACACCTAAACCTATCTCATTTGCTTACGCTTGGGGTGGAGAAC 1065
 Qy 261 LeuGlnSerGlnValLysValAlaLeuLysAlaGlySerGlySerValLeuGln 280
 Db 1066 TTAAAGTCTCAAGTCAAGTCTGCGCCCTGTGCTGTCTGGAGACAGTCTCTGCG 1125
 Qy 281 GlyValGluAlaAspLeuTyrLeuThrGlyGluMetSerHisAspThrLeuAspAla 300
 Db 1126 GGTGTGGCTGACCTTACCTCAGAGTGAATGATCCCATGATGATCTTGATCT 1185
 Qy 301 AlaSerGlnGlyLeuAsnValIleLeuGlyGlnHisSerAsnThrGluArgGlyPheLeu 320
 Db 1186 GCTTCCCAAGAAATAATGATCATCTCTGTGACACACACACACAGAGAGGCTTCT 1245
 Qy 321 SerAspLeuArgAspMetLeuAspSerHisLeuGlnAsnLysIleAsnIleLeuSer 340
 Db 1246 TCTGACCTTCAGATATCTGATCTCTGATCTGAGAAATGAATTAATCTTATCA 1305
 Qy 341 GluThrAspArgAspProLeuGlnValAla 350
 Db 1306 GAGACTGACAGGACCTCTTCAAGTGCTA 1335
 RESULT 8
 HSM805522
 LOCUS HSM805522 1353 bp mRNA linear PRI 12-JUL-2002
 DEFINITION Homo sapiens mRNA; cDNA DKFZp762L015 (from clone DKFZp762L015).
 ACCESSION AL834430
 VERSION AL834430.1 GI:21740154
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1353)
 Blecker,H., Boecher,M., Brandt,P., Mewes,H.W., Well,B. and
 Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuberger, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Brunswickweil/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp762L015) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.
 FEATURES
 Location/Qualifiers

source 1. 1353
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp762L015"
 /issue_type="melanoma (Memo cell line)"
 /clone_lib="762 (synonym: hmel2). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 polyA_signal
 polyA_site 1252
 BASE COUNT 457 a 283 c 280 g 333 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 1353
 Score: 332.00 Matches: 332
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 94.86% Indels: 0
 DB: 9 Gaps: 0
 US-09-745-506-37 (1-350) x HSM805522 (1-1353)
 Qy 19 AlaGlnSerTyrPaspAsnValGlyLeuLeuValGluProSerProProHisThrValAsn 38
 Db 2 GCTGAGACTGGGACAAATGTTGATTTACTGTGGACCAACCCACCATACCTGTAAT 61
 Qy 39 ThrLeuPheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysLys 58
 Db 62 ACACCTCTCGACCAATGACCTGACTGAGAAAGTATGAGAGAGTCTGCCAAAAGANG 121
 Qy 59 AlaAspLeuIleLeuSerTyrHisProProIlePheArgPrometLysArgIleThrTrp 78
 Db 122 GCAGACCAATCTCTCTACATCCGCTTATCTTCCGACCAATGATGATGATGATGAT 181
 Qy 79 AsnThrTyrPaspGluArgLeuValIleArgAlaLeuGlnAsnArgValGlyIleTyrSer 98
 Db 182 AACCATGGAAGAGACCGCTGATGCCGGCTCGGAGAAAGAGTGTATGATCT 241
 Qy 99 ProHisThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeu 118
 Db 242 CCTCATACAGCCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
 Qy 119 GlyAlaCysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGluGly 138
 Db 302 GAGGCTGTACTCCAGGCCCATACATCTTCCAAAGCTCCCACTACCTACAGAGGGA 361
 Qy 139 AsnHisArgValGluPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAla 158
 Db 362 AACCAACCGATGATTTACACCTTACACACCAAGACCTGACAAAGTATGCTGCA 421
 Qy 159 ValLysGlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGluGln 178
 Db 422 GTGAAGAAGAAATGACGGTGTCTGTCTCTCTTCTTCTGACAGACGTGATAGGAA 481
 Qy 179 GlnThrArgIleAsnLeuAsnGlyThrGlnLysAlaLeuMetGlnValAlaAspPheLeu 198
 Db 482 CAACACCGATTAATCTGAATGTACTCAGAAAGGCTTGATGACAGGTGATGATTTCT 541
 Qy 199 SerArgAsnLysGlnLeuTyrGlnLysThrGluIleLeuSerLeuGluLysProLeuLeu 218
 Db 542 TCCCGAACAACCACTTATTCAGAAAGACGAAATTCGTCTCTGAGAAAGCTTGTCT 601
 Qy 219 LeuHisThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuAlaThrMet 238
 Db 602 CTACATATCGGAATGAGAGCGTTATGACACATGATGATGATGATGATGATGATG 661
 Qy 239 IleAspArgIleLysArgHisLeuLysLeuSerHisTleArgLeuAlaLeuGlyValGly 258
 Db 662 ATTGATCGAATTAACCACTTAAACCTTAATCTCATATTCGTTAGCCCTTGGGTGG 721
 Qy 259 ArgThrLeuGlnSerGlnValLysValAlaLeuLysAlaGlySerGlySerSerVal 278

|||||
Db 722 AGACCTTAGAGTCTCAAGTCAAGTCTGCGCCCTGTCTGCTTCTGG3ACAGCGTT 781
Qy 279 LeuGlnGlyValGlnLysLeuLeuTyrLeuThrGlyGlnMetSerHisH13AspThrLeu 298
Db 782 CTCACAGGGTGTAGGCTGACCTTTTACCTACAGGTGAGTGTCCATCATCTGATTTG 841
Qy 299 AspAlaAlaSerGlnGlyIleAsnValIleLeuGlnGlyHisSerAsnThrGlyArgGly 318
Db 842 GATGCGCTTCCCAAGGAATAATGTATCTCTGTGACACAGCAACACTGAAACGAGGC 901
Qy 319 PheLeuSerAspLeuArgAspMetLeuAspSerHisLeuGlnAsnLysIleValIle 338
Db 902 TTTCCTTGTGACCTTGACATATGCTGATTTCTCACTTGAGAAATGAATTAATATTATC 961
Qy 339 LeuSerGlnThrAspArgAspProLeuGlnValVal 350
Db 962 CTATCAGAGCTGACAGGACCTCTTCAGGTGTA 997
RESULT 9
AF182416 1579 bp mRNA linear PRI 20-SEP-2000
LOCUS Homo sapiens MDS015 (MDS015) complete cds.
ACCESSION AF182416 GI:10197631
VERSION AF182416.1 GI:10197631
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1579)
AUTHORS Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.
TITLE Novel genes expressed in hematopoietic stem/progenitor cells from
Myelodysplastic Syndromes patient
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1579)
AUTHORS Huang, C., Qian, B., Tu, Y., Gu, W., Wang, Y., Han, Z. and Chen, Z.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, People's Republic of China
FEATURES
source
1..1579
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="hematopoietic stem/progenitor cells"
/note="Isolated from a Myelodysplastic Syndrome patient"
1..1579
/gene="MDS015"
244..1239
/gene="MDS015"
/codon_start=1
/product="MDS015"
/protein_id="AA014952.1"
/db_xref="GI:10197632"
/translation="MDLKALLSLNDFASLFAESMDWNGILVESPPTVNTFLTN
DLTEEMEVLOKKADLIISHPPIRFPKRIKLEHNGERIVATRLLENVGIYSPT
AYDAPOGVNMTAKIGIACSTSPRIHSPKAPNPTGNNRHYENVATDQDDVMSAY
KGIIDYVTSFSAKRGNEBOTRNLNCTOKALMQVDFLSRNQCLCKEITLSLEPL
LHTGMRICLTLDSEVSLATMIDRIKRLKLSHRLALGVGRTL ESQVRYVALACASG
SSVLCQVADLVLTGEMSHHDTLDAASOGINVILCHSHTNERGFLSDLRLMDLSHEN
KINIILSETRDRIQVY"
BASE COUNT 449 a 354 c 360 g 416 t
ORIGIN
Alignment Scores:
Pred. No.: 1.05e-315 Length: 1579
Score: 311.00 Matches: 350
Percent Similarity: 99.15% Conservative: 0
Best Local Similarity: 99.15% Mismatches: 0
Query Match: 88.86% Indels: 3
DB: 9 Gaps: 0

US-09-745-506-37 (1-350) x AF182416 (1-1579)
Qy 1 MetAspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20
Db 244 ATGGATTTGAAGCTCTCCCTTTCTCTTGAATGATCTTCATCCCTCTGTTGGTGGAG 303
Qy 21 SerTyrAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
Db 304 AGTTGGACAAATGTTGATTAATGTTGTTGAAACCAAGCCACACACTGTAATTAACATC 363
Qy 41 PheLeuThrAsnAspLeuThrGlnGluValMetGlnGluValLeuGlnLysAlaAsp 60
Db 364 TTCCCTGACCAATGACCTGACTGAGGAAGTATGAGAGAGGTGCTGCAAAAGAGCGAGAC 423
Qy 61 LeuIleLeuSerTyrHisProProIlePheArgProMetLysArgIle-ThrTyrAsn 80
Db 424 CTCATTCCTCCCTACACCTCCGCTATCTCCGACCCCATGGAAGCGCATTAACCTGGAACAC 483
Qy 80 rTTP-Lys-GluArgLeuValIleArgAlaLeuGlnAsnArgValGlyIleTyrSerPro 99
Db 484 ATGGGAAGGAGAGCGCTGTGATCCGGCTCTGGAGACAGAGTGGTATCTACTCTCT 543
Qy 100 HisThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnThrPheAlaLysGlyLeuGly 119
Db 544 CATACAGCCTATGATGTGCGCCCGCCAGGGGCTCACACAACTGGTTGGCTAAAGGGCTTGA 603
Qy 120 AlaCysThrSerArgProIleHisProSerLysAlaProAsnTyrProThrGlnGlyAsn 139
Db 604 GCTTGTATACCTCCAGGCCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGGGAAC 663
Qy 140 HisArgValGlnPheAsnValAsnTyrThrGlnAspLeuAspLysValMetSerAlaVal 159
Db 664 CACCGATACAAATTCACGTTAACTACACCAAGACCTGGACAAAGCAATGCTGACGTG 723
Qy 160 LysGlyIleAspGlyValSerValThrSerPheSerAlaArgThrGlyAsnGlnGln 179
Db 724 AAGGAATTCACGTTGTTCTGACATCTTTTCTGCTGAGACTGTAATGAGAAACA 783
Qy 180 ThrArgIleAsnLeuAsnCysThrGlnLysAlaLeuMetGlnValValAspPheLeuSer 199
Db 784 ACACGGATTAATCTGAATTTACTCAGAAAGCTTTGATGAGGTGAGATTTTCTTCC 843
Qy 200 ArgAsnLysGlnLeuTyrGlnLysThrGlnIleLeuSerLeuGlnLysProLeuLeu 219
Db 844 CGGACAAACAACTTATTCAGAAAGCGAAATTTCTGACTGGAGAAAGCTTTGCTTCTA 903
Qy 220 HisThrGlyMetGlyArgLeuGlyCysThrLeuAspLysSerValSerLeuAlaThrMetIle 239
Db 904 CATACGTGAATGAGGACGTTATGCACACTGAGTAATCTGTCTCCCGCAACCATGATT 963
Qy 240 AspArgIleLysArgHisLeuLysLysLeuSerHisIleArgLeuAlaLeuGlyValArg 259
Db 964 GATCGAATAAAGACACCTTAACATTCATTCATTCGTTACCCCTTGGGGTGGGAGA 1023
Qy 260 ThrLeuGlnSerGlnValLysValAlaIleLeuGlnLysGlySerGlySerValLeu 279
Db 1024 ACCTTAGAGCTCAAGTCAAAAGTCTGCGCCCTGTCTGCTGTTCTGGAGACACGTTCTG 1083
Qy 280 GlnGlyValGlnLysLeuTyrLeuThrGlyGlnMetSerHisAspThrLeuAsp 299
Db 1084 CAGGGGTGTAGGCTGACCTTACCTACACAGTGAAGATGCCCATCATGATTAATTGGAT 1143
Qy 300 AlaAlaSerGlnGlyIleAsnValIleLeuGlnGlnHisSerAsnThrGlnArgGlyPhe 319
Db 1144 GCTGCTTCCCAAGGAATAATGTATCTCTGTGACACAGCAACACTGAAACAGCGTTT 1203
Qy 320 LeuSerAspLeuArgAspMetLeuAspSerHisLeuGlnLysAsnLysIleValIleLeu 339
Db 1204 CTTCGTGACCTTGAGATATGCTGATTCCTCACTTGAGAAATGAATTAATATTATTCCTA 1263
Qy 340 SerGlnThrAspArgAspProLeuGlnValVal 350
|||||

Db 1264 TCAGAGCTGACGAGGACCTCTTCAGCTGTA 1296

RESULT 10

LOCUS BD149184 796 bp DNA linear PAT 17-JAN-2003

DEFINITION BD149184 primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD149184.1 GI:27854942

VERSION JP 2002191363-A/4027.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 796) Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,T., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 4027 09-JUL-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/4027

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU

PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,

PI KEIICHI MAGAI,TETSUJI OTSUKI

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC

10,

PC C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH key

Location/Qualifiers

FT source 1..796

FT /organism='Homo sapiens (human)'

1..796

/mol_type='genomic DNA'

/db_xref='taxon:9606'

BASE COUNT 192 a 204 c 202 g 195 t 3 others

ORIGIN

Alignment Scores:

Pred. NO.: 7.17e-118 Length: 796

Score: 122.00 Matches: 122

Percent Similarity: 100.00% Conservative: 0

Best local Similarity: 100.00% Mismatches: 0

Query Match: 34.86% Indels: 0

DB: Gaps: 0

US-09-745-506-37 (1-350) x BD149184 (1-796)

Oy 1 MetaspLeuLysAlaLeuLeuSerSerLeuAsnAspPheAlaSerLeuSerPheAlaGlu 20

Db 271 ATGATTTGAAAGCGCTCTTCTTCCTTGAATGACTTGCATCCCTCGTTGCTGAG 330

Oy 21 SerTrrAspAsnValGlyLeuLeuValGluProSerProProHisThValAsnThriu 40

Db 331 AGTTGGACAAATGTTGATCTACTGTTGGAACCAAGCCACACATACATCTAAATCACTC 390

Oy 41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysLysAlaAsp 60

Db 391 TTCTCGAACCAATGACTGACTGAGAACTGATGAGAGGTCTGCAAAAGAGGCAAC 450

Oy 61 LeuLeuSerTrrHisProProLeuPheArgProMetLysArgIleThrTrrAsnThr 80

Db 451 CTCATCTCTCTACCATCGGCTTATCTTCGACCAATGAAAGCAATACCTGGAGACAC 510

Oy 81 TrrLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTrrSerProHis 100

Db 511 TGGAAAGAGCGCGCTGCTGATCCGGGCTCTGGAGAACAGAGTGGTATCTACTCTCAT 570

Oy 101 ThrAlaTrrAspAlaAlaProGlnGlyValAsnAsnTrrPheAlaLysGlyLeuGlyAla 120

Db 571 ACAGCCTATATGCTGCGCCCAAGCGCTACACACATGCTTGGCTAAAGGCTTGAGCT 630

Oy 121 CysThr 122

Db 631 TGTACC 636

RESULT 11

AC037455/c

LOCUS AC037455 170586 bp DNA linear HTG 30-AUG-2001

DEFINITION Homo sapiens chromosome 02 clone RP11-422L5, WORKING DRAFT

ACCESSION AC037455

VERSION AC037455.5 GI:9887641

KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS-CANCELLED.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 170586) Smith,D.R.

REFERENCE Smith,D.R.

AUTHORS Genome Therapeutics Corporation Sequencing Center: Human Genome

TITLE Sequence Data

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 170586)

TITLE Smith,D.R.

JOURNAL Direct Submission

COMMENT Submitted (09-APR-2000) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA

On Aug 24, 2000 this sequence version replaced g1:8569072.

Center: Genome Therapeutics Corporation

Center code: GTC

Web site: <http://www.genomecorp.com/>

Contact: gtc-seqcenter@genomecorp.com

Project Information

Center project name: hg215

Summary Statistics

Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 990315

Consensus quality: 147001 bases at least Q40

Consensus quality: 157396 bases at least Q30

Consensus quality: 159713 bases at least Q20

Insert size: 168303; sum-of-contigs

Quality coverage: 3.8x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 25 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1171: contig of 1171 bp in length

1172 1271: gap of unknown length

1272 2319: contig of 1048 bp in length

2320 2419: gap of unknown length

2420 3793: contig of 1374 bp in length

3794 3893: gap of unknown length

3894 5414: contig of 1521 bp in length

5415 5514: gap of unknown length

5515 8773: contig of 3259 bp in length

8774 8873: gap of unknown length

8874 10904: contig of 2031 bp in length

10905 11004: gap of unknown length

11005 13204: contig of 2200 bp in length

13205 13304: gap of unknown length

13305 16401: contig of 3037 bp in length

16402 16501: gap of unknown length

16502 18327: contig of 1726 bp in length

*	18327:	gap of unknown length
*	18328	
*	22402	22501: contig of 4074 bp in length
*	22502	22501: gap of unknown length
*	25177:	contig of 2676 bp in length
*	25178	25277: gap of unknown length
*	25278	28992: contig of 4615 bp in length
*	28983	29993: gap of unknown length
*	29993	33457: contig of 3465 bp in length
*	33458	33557: gap of unknown length
*	33558	37405: contig of 3848 bp in length
*	37406	37505: gap of unknown length
*	37506	41734: contig of 4229 bp in length
*	41735	41834: gap of unknown length
*	41835	48999: contig of 7165 bp in length
*	49000	49099: gap of unknown length
*	49100	56076: contig of 6977 bp in length
*	56077	56176: gap of unknown length
*	56177	63604: contig of 7428 bp in length
*	63605	63704: gap of unknown length
*	63705	70821: contig of 7117 bp in length
*	70822	70921: gap of unknown length
*	70922	80138: contig of 9217 bp in length
*	80139	80238: gap of unknown length
*	80239	88780: contig of 8542 bp in length
*	88781	88880: gap of unknown length
*	88881	104642: contig of 13762 bp in length
*	104643	104742: gap of unknown length
*	104743	124656: contig of 19914 bp in length
*	124657	124756: gap of unknown length
*	124757	143891: contig of 19135 bp in length
*	143892	145991: gap of unknown length
*	143992	170586: contig of 26595 bp in length

location/Qualifiers	
source	1. 170586
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="02"
	/clone="RP11-422L5"
	/clone_11b="RPCI-11"
	1. 1171
misc_feature	/note="assembly_name:Contig1"
misc_feature	1272..2319
	/note="assembly_name:Contig13"
misc_feature	2430..3793
	/note="assembly_name:Contig17"
misc_feature	3894..5414
	/note="assembly_name:Contig18"
misc_feature	5515..8773
	/note="assembly_name:Contig19"
misc_feature	8874..10904
	/note="assembly_name:Contig20
	clone_end:"7"
misc_feature	11005..11304
	/note="assembly_name:Contig21"
misc_feature	13305..16301
	/note="assembly_name:Contig22"
misc_feature	16502..18327
	/note="assembly_name:Contig23"
misc_feature	18328..22401
	/note="assembly_name:Contig24"
misc_feature	22502..25177
	/note="assembly_name:Contig25"
misc_feature	25278..29892
	/note="assembly_name:Contig26"
misc_feature	29993..33357
	/note="assembly_name:Contig27"
misc_feature	3358..37405
	/note="assembly_name:Contig28"
misc_feature	37506..41134
	/note="assembly_name:Contig29"
misc_feature	41835..46899
	/note="assembly_name:Contig30"

	misc_feature	49100..56076	/note="assembly_name:Contlg31"	
	misc_feature	56177..63604	/clone_end=SP6	
	misc_feature	63705..70821	/note="assembly_name:Contlg32"	
	misc_feature	70922..80138	/note="assembly_name:Contlg33"	
	misc_feature	80233..88780	/note="assembly_name:Contlg34"	
	misc_feature	88881..104642	/note="assembly_name:Contlg35"	
	misc_feature	104743..124656	/note="assembly_name:Contlg36"	
	misc_feature	124757..143891	/note="assembly_name:Contlg37"	
	misc_feature	143992..170586	/note="assembly_name:Contlg38"	
	misc_feature		/note="assembly_name:Contlg39"	
BASE COUNT	50417 a	34173 c	33230 g	50259 t
ORIGIN			2447 others	
 Alignment Scores:				
Pred. No.:	9.26e-113	Length:	170586	
Score:	119.00	Matches:	119	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	34.00%	Indels:	0	
DB:	2	Gaps:	0	
 US-09-745-506-37 (1-350) x AC037455 (1-170586)				
OY	1	MetaSpleLeuValLeuLeuSerSerLeuAaAspPheAlaSerLeuSerPheAlaGlu	20	
Db	140069	ATGGATTGGAAGGCCTCCTTCTTCCTTGAAATGACTTTGCATCCCTCGTTGGCTGAG	140010	
OY	21	SerTrpAspAsnValGlyLeuLeuValGluProSerProProHISThrValAsnThrLeu	40	
Db	140009	AGTTGGACCAATGTGGATTACTGCGTGAACAACCACCATCATCTGTAATATACACTC	139950	
OY	41	PheLeuThrAspAspLeuThrGluGluValMetGluGluValLeuGlnLysAlaAsp	60	
Db	139949	TTCCTGCACCAATGACCTGACTGAGAAGATGAGAGAGGCTGCAAAAAGAAAGCGAC	139890	
OY	61	LeuIleLeuSerTyHisProProIlePheArgProMetLysArgIleThrTrpAsnThr	80	
Db	139889	CTCATTCCTCCTTACCATCCGGCTATCTCCGACCATGAGGCCATTAACCTGGAAACACA	139830	
OY	81	TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTySerProHis	100	
Db	139839	TGGAGGAGAGCCCTGGTGATCGGGCTGGAACAAGAGTGGTATCTACTCTCCCAT	139770	
OY	101	ThrAlaTyRAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGly	119	
Db	139769	ACAGCCTATGATGCTCGCCGCCACAGGCGTCAACAACATGCTTGCGTTAAGGCGTTGGT	139713	
 RESULT 12				
LOCUS	AC005037/C	190508 bp	DNA	linear PRI 30-SEP-2000
DEFINITION	Homo sapiens BAC clone RP11-46M7 from 2, complete sequence.			
ACCESSION	AC005037			
VERSION	AC005037.2	GI:4827310		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 190508)			
JOURNAL	Toward a complete human genome sequence			
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)			
PUBMED	99063792			
	9847074			

REFERENCE 2 (bases 1 to 190508)
 AUTHORS Abbott, A. and Le, T.
 TITLE The sequence of Homo sapiens BAC clone RP11-469M7
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 190508)
 AUTHORS Waterston, R. H.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 190508)
 AUTHORS Waterston, R. H.
 TITLE Direct Submission
 JOURNAL Submitted (14-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 190508)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On May 14, 1999 this sequence version replaced gi:3309089.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0469M07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catanesi, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-1338; the clone sequenced to the right is RP11-91M5. Actual start of this clone is at base position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.

FEATURES

source

1. 190508
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="RP11-469M7"
 /clone_id="RP11-11"

repeat_region	155..222
repeat_region	/rpt_family="L2"
repeat_region	982..1299
repeat_region	/rpt_family="Alu"
repeat_region	1502..1533
repeat_region	/rpt_family="(CA)n"
repeat_region	1535..1823
repeat_region	/rpt_family="Alu"
repeat_region	2005..2308
repeat_region	/rpt_family="Alu"
repeat_region	2393..2562
repeat_region	/rpt_family="L1"
repeat_region	3284..3315
repeat_region	/rpt_family="Alu"
repeat_region	3316..3338
repeat_region	/rpt_family="(CANA)n"
repeat_region	3339..3621
repeat_region	/rpt_family="Alu"
repeat_region	3650..3743
repeat_region	/rpt_family="AT-rich"
repeat_region	4353..4664
repeat_region	/rpt_family="Alu"
repeat_region	4727..5018
repeat_region	/rpt_family="Alu"
repeat_region	5545..5585
repeat_region	/rpt_family="(CA)n"
repeat_region	5586..5873
repeat_region	/rpt_family="Alu"
repeat_region	5876..6007
repeat_region	/rpt_family="Alu"
repeat_region	6327..6713
repeat_region	/rpt_family="MER21-g"
repeat_region	6924..7294
repeat_region	/rpt_family="MER2_type"
repeat_region	8027..8049
repeat_region	/rpt_family="AT-rich"
repeat_region	8050..8354
repeat_region	/rpt_family="Alu"
repeat_region	8355..8661
repeat_region	/rpt_family="Alu"
repeat_region	8356..8684
repeat_region	/rpt_family="AT-rich"
repeat_region	9450..9752
repeat_region	/rpt_family="Alu"
repeat_region	10371..10680
repeat_region	/rpt_family="Alu"
repeat_region	10762..11038
repeat_region	/rpt_family="Alu"
repeat_region	11039..11109
repeat_region	/rpt_family="Alu"
repeat_region	11138..11410
repeat_region	/rpt_family="Alu"
repeat_region	11446..11663
repeat_region	/rpt_family="L1"
repeat_region	11664..11959
repeat_region	/rpt_family="Alu"
repeat_region	11960..12189
repeat_region	/rpt_family="L1"
repeat_region	12190..12334
repeat_region	/rpt_family="Alu"
repeat_region	12335..12360
repeat_region	/rpt_family="L1"
repeat_region	12464..12514
repeat_region	/rpt_family="L1"
repeat_region	12515..12822
repeat_region	/rpt_family="Alu"
repeat_region	12823..12862
repeat_region	/rpt_family="L1"
repeat_region	12857..13096
repeat_region	/rpt_family="L1"
repeat_region	13101..13330
repeat_region	/rpt_family="MALR"
repeat_region	13331..13634

```

repeat_region      /rpt_family="Alu"
                    13635..13692
repeat_region      /rpt_family="MELR"
                    13704..13990
repeat_region      /rpt_family="L1"
                    14082..14387
repeat_region      /rpt_family="Alu"
                    14402..14704
repeat_region      /rpt_family="Alu"
                    14797..15095
repeat_region      /rpt_family="Alu"
                    15162..15196
repeat_region      /rpt_family="Mariner"
                    15606..15681
repeat_region      /rpt_family="L1"
                    15691..15827
repeat_region      /rpt_family="MER1_type"
                    15861..16189
repeat_region      /rpt_family="L2"
                    16426..16446
repeat_region      /rpt_family="(CAAT)n"
                    16447..16730
repeat_region      /rpt_family="Alu"
                    17072..17179
repeat_region      /rpt_family="L2"
                    17212..17504
repeat_region      /rpt_family="Alu"
                    17869..18139
repeat_region      /rpt_family="Alu"
                    18168..18464
repeat_region      /rpt_family="Alu"
                    18606..19155
repeat_region      /rpt_family="MELR"
                    19454..19767
repeat_region      /rpt_family="Alu"
                    19807..19949
repeat_region      /rpt_family="MER1_type"
                    19950..20247
repeat_region      /rpt_family="Alu"

```

```

Alignment Scores:
Pred. No.:          1,02e-112      Length:      190508
Score:              119.00          Matches:      119
Percent Similarity: 100.00%         Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:        34.00%          Indels:      0
DB:                  9              Gaps:         0

```

US-09-745-506-37 (1-350) x AC005037 (1-190508)

```

OY      1 MetaspLeuLYAlaLeuSerSerLeuAanAspPheAlaSerLeuSerPheAlaGu 20
Db      122461 ATGGATTTCAGAGCTCTCTTCTTCTTCTTGAATGACTTTCCTCTGTTGCTGAG 122402
OY      21 SerTPAspAsnValGlyLeuLeuValGluProSerProProHisThrValAsnThrLeu 40
Db      122401 AGTTGGGACAAATGTTGATTACTGTTGGTGAACCAAGCCACACATCTGTAATACATC 122342
OY      41 PheLeuThrAsnAspLeuThrGluGluValMetGluGluValLeuGlnLysLysAlaasp 60
Db      122341 TTCCCTGACCAATGACTGCTGACTGAGCAAGATGATGAGAGAGTCTGCCAAAGAAGCAGAC 122282
OY      61 LeuileuSerTyrHisProProIlePheArgProMetLysArgIleThrTrpAsnThr 80
Db      122281 CTCATTCTCTCTACATCGCTTCTTCTTCTTCCGACCATCAAGCGCATTAACCTGGAACACA 122222
OY      81 TrpLysGluArgLeuValIleArgAlaLeuGluAsnArgValGlyIleTyrSerProHis 100
Db      122221 TGGAGGAGACGCGCTGGTGTATCGGGGCTCGGAGAAAGAGTGGTATCTTCTCTCTCAT 122162
OY      101 ThrAlaTyrAspAlaAlaProGlnGlyValAsnAsnTrpLeuAlaLysGlyLeuGly 119
Db      122161 ACAGCTATGATGCTGCGCCCGACAGGCGTCACAACTGCTGCTAAAGGCTTGCT 122105

```

```

RESULT 13
AC093681
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

COMMENT

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wustl.wustl.edu
 Project Information
 Center project name: H_NH0663N02

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 26 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      1912: contig of 1912 bp in length
*      1913      2012: gap of unknown length
*      2013      3539: contig of 1527 bp in length
*      3540      3639: gap of unknown length
*      3640      6127: contig of 2488 bp in length
*      6128      6227: gap of unknown length
*      6228      9244: contig of 3017 bp in length
*      9245      9345: gap of unknown length
*      9345      13288: contig of 3944 bp in length
*      13289      13388: gap of unknown length
*      13389      16301: contig of 2913 bp in length
*      16302      16401: gap of unknown length
*      16401      20096: contig of 3695 bp in length
*      20097      20196: gap of unknown length
*      20197      24467: contig of 4271 bp in length
*      24468      24567: gap of unknown length
*      24568      28680: contig of 4113 bp in length
*      28681      28780: gap of unknown length
*      28781      32494: contig of 3714 bp in length
*      32495      32594: gap of unknown length
*      32595      35726: contig of 3032 bp in length
*      35727      35726: gap of unknown length
*      35727      40050: contig of 4324 bp in length
*      40051      40150: gap of unknown length
*      40151      43881: contig of 3731 bp in length
*      43882      43981: gap of unknown length
*      43982      50194: contig of 6213 bp in length
*      50195      50294: gap of unknown length
*      50295      57454: contig of 7160 bp in length
*      57455      57554: gap of unknown length
*      57555      64660: contig of 7106 bp in length
*      64661      64760: gap of unknown length

```

	*	64761	72159:	contig of 7399 bp in length
	*	72160	72359:	gap of unknown length
	*	72260	79835:	contig of 7576 bp in length
	*	79836	79935:	gap of unknown length
	*	79936	90863:	contig of 10928 bp in length
	*	90864	90963:	gap of unknown length
	*	99400	99400:	contig of 8437 bp in length
	*	99501	99500:	gap of unknown length
	*	114190	114190:	contig of 14690 bp in length
	*	114191	114290:	gap of unknown length
	*	114291	127822:	contig of 13352 bp in length
	*	127823	127922:	gap of unknown length
	*	127923	143474:	contig of 15552 bp in length
	*	143475	143574:	gap of unknown length
	*	143575	170444:	contig of 26870 bp in length
	*	170445	170544:	gap of unknown length
	*	170545	195551:	contig of 25107 bp in length
	*	195552	195751:	gap of unknown length
	*	195752	198250:	contig of 2499 bp in length.
FEATURES			Location/Qualifiers	
source			1..198250	
			/organism="Homo sapiens"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9606"	
			/chromosome="4"	
			/clone="RP11-663N2"	
			1..1912	
misc_feature			/note="assembly_name:Contig10"	
			2013..3539	
misc_feature			/note="assembly_name:Contig11"	
			3640..6127	
misc_feature			/note="assembly_name:Contig12"	
			6228..9244	
misc_feature			/note="assembly_name:Contig13"	
			9345..13288	
misc_feature			/note="assembly_name:contig14"	
			13389..16301	
misc_feature			/note="assembly_name:Contig15"	
			16402..20096	
misc_feature			/note="assembly_name:Contig16"	
			20197..24467	
misc_feature			/note="assembly_name:Contig17"	
			24568..28680	
misc_feature			/note="assembly_name:Contig18"	
			28781..32494	
misc_feature			/note="assembly_name:Contig19"	
			32595..35626	
misc_feature			/note="assembly_name:Contig20"	
			35727..40050	
misc_feature			/note="assembly_name:Contig21"	
			40151..43881	
misc_feature			/note="assembly_name:Contig22"	
			43982..50194	
misc_feature			/note="assembly_name:Contig23"	
			50295..57454	
misc_feature			/note="assembly_name:Contig24"	
			57555..64660	
misc_feature			/note="assembly_name:Contig25"	
			64761..72159	
misc_feature			/note="assembly_name:Contig26"	
			72260..79835	
misc_feature			/note="assembly_name:Contig27"	
			79936..90863	
misc_feature			/note="assembly_name:Contig28"	
			90964..99400	
misc_feature			/note="assembly_name:Contig29"	
			99501..114190	
misc_feature			/note="assembly_name:Contig30"	
			114291..127822	
misc_feature			/note="assembly_name:Contig31"	
			127923..143474	
misc_feature			/note="assembly_name:Contig32"	
			143575..170444	

	misc_feature	/note="assembly_name:contig93" 170545..195651 /note="assembly_name:contig93"
	misc_feature	/note="assembly_name:contig93" 195752..198250 /note="assembly_name:contig93"
BASE COUNT	56011 a	42355 c 41444 g 55928 t 2512 others
ORIGIN		
<hr/>		
Alignment Scores:		
Pred. No.::	1_05e-112	Length: 198250
Score:	119.00	Matches: 119
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	34.00%	Indels: 0
DB:	2	Gaps: 0
<hr/>		
US-09-745-506-37 (1-350) x AC093681 (1-198250)		
QY	I MetAspleuysAlaleuleusSerSeriEuasnsPhealAserLeusErphealaGu 20	
Dd	14147 ATGGAATTGAAGCCTCCCTTTCCTCGTAATGACTTTGCATCCCTCYGTGGTGAG 14206	
QY	21 SetTrpaspsnAlaglyleuVeuValguProSePrroPhoishrThralasnThrIeu 40	
Dd	14207 AGTGGACAAATGTGGATTACTGGGGAAACAAGCCCACCATACTGTAATACACTC 14266	
QY	41 PheueurhzansApneuThrgluGlualMetgluglualLenglnhsyrsylaaSp 60	
Dd	14267 TTCTGCACCAATGCCGACTGAGAAGTAGCATGAGAGGTGCTGCAAAAAGAACGACAG 14326	
QY	61 LeuilEeuserryrhSPpropolliepharygPromellysarglylerhtrrpaSnthr 80	
Dd	14327 CTATTTCCTCACCACTCCGCTCACTTCGACCACTGAAGGCGCTMAACCTGGAACACA 14386	
QY	81 TrplsGLuarqleunvalilearGalalaleuguUasnarGaylaGlyrlerySerprohis 100	
Dd	14387 TGGAGAGAGCGCCTGGTGATCCGGGCGCTGGANAACAGATGGCTATCACTCTCAT 14446	
QY	101 ThrAlatryrsapAlaalaproginglyyalnasnsnrTripeualaalysglyencly 119	
Dd	14447 ACAACCTATGATGTGCGCCCCAGGGGCTCAACACTGGTGGCTAAAGGGCTTGGT 14503	
<hr/>		
RESULT 14		
BD049005		
LOCUS	BD049005	249 bp DNA linear PAT 27-AUG-2002
DEFINITION	Sequence tag and encoded human protein.	
ACCESSION	BD049005	
VERSION	BD049005.1	GI:22590747
KEYWORDS	JP 2001269182-A/25251.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
	1 (bases 1 to 249)	
REFERENCE	Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.	
AUTHORS	Sequence tag and encoded human protein	
TITLE	Patent: JP 2001269182-A 25251 02-OCT-2001.	
JOURNAL	GENSET	
COMMENT	OS Homo sapiens (human) PN JP 2001269182-A/25251 PD 02-OCT-2001 PE 24-FEB-2000 JP 2000118773 PR 26-FEB-1999 US 60/122487 PI JEAN BABUTIST DUMAS MILNE EDWARDS.EIMERIC DUCLAIR.JEAN YVES JORDAN PC C12N15/09.C07K14/435.C07K16/18.C12N1/15.C12N1/19.C12M1/21.PC C12N5/10. PC C12P21/02.C12P21/08.C12O1/68//G06F17/30.C12N15/00.C12N5/00.PC G06F15/40	
FEATURES	key	Location/Qualifiers.
CDS	FH	location/Qualifiers source 1..249

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT      62 a      65 c      57 g      63 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      8.53e-64      Length:      249
Score:          70.00      Matches:      70
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    20.00%      Indels:      0
DB:            6      Gaps:      0

US-09-745-506-37 (1-350) x BD049005 (1-249)

OY      10 LeuansapPheal1SerleuserPhealagIuserTPaspasValcIyleuVal 29
      37 TTGAATGACTTGCCTCCCTCTGCTGAGAGCTTGAGCAATGCTGATTAAGTGGTG 96
      30 GUProserProH1sthrValasnThrlaupheleuthrAsnAspleuthrGlugu 49
      97 GACCAACCCACCAACATACCTAATACACTCTTCTGACCAATGACTGACCTGAGAA 156
OY      50 ValMetgluValleuGlInLysLysAlaAspleuileuserTyrlHisProProle 69
      157 GTGATGAGAGAGAGCTCTGCAAAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216
OY      70 PheaIgrPrometLysArgIleThrtPasn 79
      217 TTCGACCCATGACGATACCTGACGAGAC 246

RESULT 15
AC130779      231600 bp      DNA      linear      HTG 09-NOV-2002
LOCUS      Rattus norvegicus clone CH230-200A21, WORKING DRAFT SEQUENCE. 3
DEFINITION      AC130779
AC130779.3 GI:24818669
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FILLTOP.
KEYWORDS      Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM      Rattus
      1 (bases 1 to 231600)
      Muzny,D,Marle, Metzker,M, Lee, Abramzon,S, Adams,C, Alder,J,
      Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguliano,D,
      Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
      Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
      Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
      Bryant,N, Bunay,C, Burch,P, Butrell,K, Calderon,E,
      Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Center,A,
      Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
      Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
      Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
      Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
      Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
      Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
      Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
      Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
      Gheorgheorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
      Gunaratne,P, Haaland,W, Hamli,C, Hamilton,C, Hamilton,K,
      Harvey,Y, Havlik,P, Hawes,A, Henderson,N, Hernandez,U,
      Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
      Hollins,B, Howells,S, Huiyuk,S, Hume,J, Idlebird,D, Jackson,A,
      Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
      Karpaty,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
      Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
      Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
      Lounshewa,L, Louissege,H, Lozado,R,U, Lu,X, Ma,J,
      Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A,
      Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,

```

```

Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Parks,K,
Nwakoilemen,O, Okwunonu,G, Olarinuagun,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C,
Plopper,F, Polidexter,A, Popovic,D, Prims,E, Pu,L,
Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rivers,C, Rodkey,T, Rojas,A, Rose,K, Rose,R, Ruiz,S,U,
Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D,
Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,R, Sosa,D,
Steinle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umani,K,
Valas,R, Vera,Y, Villaseca,D, Waldron,L, Walker,B, Wang,J,
Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
Williams,G, Willson,R, Wleczek,R, Wooden,H, Worley,K,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,D, Yoon,L, Yoon,V,
Yu,F, Zhang,Y, Zhou,X, Zhou,S, Zhou,S, Dunn,D, Von
Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
Weinstock,G, and Gibbs,R.A.
Direct Submission

Unpublished
2 (bases 1 to 231600)
Rat Genome Sequencing Consortium.
Submitted (14-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 231600)
Rat Genome Sequencing Consortium.
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced g1:22857107.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rae/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOCJ
Center clone name: CH230-200A21
----- Summary Statistics
Assembly program: Phrap; version: 0.990329
Consensus quality: 218352 bases at least Q40
Consensus quality: 221053 bases at least Q30
Consensus quality: 223057 bases at least Q20
Estimated insert size: 223361; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```


* as soon as it is available and the accession number will
* be preserved.

1 229061: contig of 229061 bp in length
* 229062 228161: gap of unknown length
* 229162 230276: contig of 1115 bp in length
* 230277 230376: gap of unknown length
* 230377 231600: contig of 1224 bp in length.

FEATURES
source
1. 231600
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-200A21"
1. 1229
/note="wgs_contig"
misc_feature 222785..224838
/note="wgs_contig"
misc_feature 225945..227046
/note="wgs_contig"
misc_feature 227515..229061
/note="wgs_contig"

BASE COUNT 65457 a 49366 c 47945 g 61775 t 7057 others
ORIGIN

Alignment Scores:
Pred. No.: 1.79e-43 Length: 231600
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 15.14% Indels: 0
DB: 2 Gaps: 0

US-09-745-506-37 (1-350) x AC130779 (1-231600)

QY 1 Metasp1euysAlaleu1euSerSerleu5na5ppheal1aserleuSerPheal1glu 20
Db 131105 ATGATCTGAAGGCTCTCTCTCTGATGACTTGGCTCCCTCTCTCATTTGCTGAG 131164
QY 21 SerTrpasp1naVal1gl1eu1euVal1gluProSerProPH1sthrVal1asnThr1eu 40
Db 131165 AGCTGGGACATGTGGGGTCTGCTGTGAGCCAGCCCGCCCATACTGTAATACACTC 131224
QY 41 Phe1euThr1asn1p1euThr1glu1gluVal1Met1glu1glu 53
Db 131225 TTCTTGACCATGACTGACGAGAGAGGTGATGAGAGAG 131263

Search completed: August 23, 2003, 21:44:25
Job time : 3126 secs